UNITED STATES DISTRICT COURT
DISTRICT OF MASSACHUSETTS

CIVIL ACTION NO. 08-11132-GAO

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
Plaintiff,

v.

AFFYMETRIX, INC.,
Defendant.

CLAIM CONSTRUCTION ORDER
September 4, 2012

O'TOOLE, D.J.

The plaintiff, Massachusetts Institute of Technology, claims that the defendant, Affymetrix, Inc., infringes United States Patent No. 6,703,228 (the "'228 Patent") entitled "Methods and Products Related to Genotyping and DNA Analysis." The parties dispute the meaning of certain terms in the patent's claims and have moved the Court to determine their proper construction in accordance with Markman v. Westview Instruments, Inc., 517 U.S. 370 (1996). The Court construes those terms as set forth herein.

The invention claimed in the '228 Patent concerns "genotyping methods involving detection of one or more single nucleotide polymorphisms (SNPs) in a reduced complexity genome (RCG) prepared from the genome of a subject." Col. 10:13-16.[1] A useful purpose of the invention is said to be improved ability to identify genetic markers for particular diseases.

In construing terms in patent claims, a court must assign to a claim term "the meaning that the term would have to a person of ordinary skill in the art in question at the time of the

[1] References to specific language in the patent are to column and line. Column 10, lines 13-16 is rendered Col. 10:13-16.

invention." Phillips v. AWH Corp., 415 F.3d 1303, 1313 (Fed. Cir. 2005) (en banc). Intrinsic evidence, which includes the patent, the claims, the specification, and the prosecution history, "is the most significant source of the legally operative meaning of disputed claim language." Vitronics Corp. v. Conceptronic, Inc., 90 F.3d 1576, 1582 (Fed. Cir. 1996).

A court may also at times consider extrinsic evidence such as dictionaries, treatises, and expert testimony. Vitronics, 90 F.3d at 1584. However, "[t]he court turns to extrinsic evidence only when the intrinsic evidence is insufficient to establish the clear meaning of the asserted claim." Zodiac Pool Care, Inc. v. Hoffinger Indus., Inc., 206 F.3d 1408, 1414 (Fed. Cir. 2000); accord C.R. Bard, Inc. v. U.S. Surgical Corp., 388 F.3d 858, 862 (Fed. Cir. 2004) ("[E]vidence extrinsic to the patent document "can shed useful light on the relevant art", but is less significant than the intrinsic record in determining the legally operative meaning of disputed claim language.") (internal citations and quotations omitted); Intel Corp. v. VIA Techs., Inc., 319 F.3d 1357, 1367 (Fed. Cir. 2003) ("When an analysis of intrinsic evidence resolves any ambiguity in a disputed claim term, it is improper to rely on extrinsic evidence to contradict the meaning so ascertained.").

A. Claim 1

Claim 1 of the '228 Patent (as corrected) [2] reads:

> A method for detecting the presence or absence of a single nucleotide polymorphism (SNP) allele in a genomic DNA sample, the method comprising:
>
> preparing a reduced complexity genome (RCG) from the genomic DNA sample, wherein the RCG is a randomly primed PCR-derived RCG, and
>
> analyzing the RCG directly by hybridization for the presence or absence of a SNP allele.

[2] The language in the original patent was corrected in a certificate of correction. (See Affymetrix Claim Construction Presentation 8; Joint Claim Construction Chart 2-5 (dkt. no. 109-2).)

1. Preamble: "method for detecting the presence or absence of a single nucleotide polymorphism (SNP) allele in a genomic DNA sample"

The parties dispute whether the claim's preamble limits the claim.

A preamble's purpose typically is "to give context for what is being described in the body of the claim." Symantec Corp. v. Computer Assocs. Int'l Inc., 522 F.3d 1279, 1288 (Fed. Cir. 2008). "A preamble is not limiting . . . where a patentee defines a structurally complete invention in the claim body and uses the preamble only to state a purpose or intended use for the invention." Catalina Mktg. Int'l v. Coolsavings.com, Inc., 289 F.3d 801, 808 (Fed. Cir. 2002) (internal citations and quotations omitted). If a preamble "is reasonably susceptible to being construed to be merely duplicative of the limitations in the body of the claim (and was not clearly added to overcome a rejection), we do not construe it to be a separate limitation." Symantec, 522 F.3d at 1288-89. However, a preamble may limit a claim "if it recites essential structure or steps, or if it is necessary to give life, meaning, and vitality to the claim." Id. at 1288 (internal citations and quotations omitted). "Generally, the preamble does not limit the claims." Allen Eng'g Corp. v. Bartell Indus., Inc., 299 F.3d 1336, 1346 (Fed. Cir. 2002).

In the present case, the preamble simply states the invention's purpose, namely, detecting a SNP allele of interest in a genomic DNA sample. The body of the claim then explains the method in two steps ("preparing" and "analyzing"). The preamble does not specify or define; it summarizes. Affymetrix argues that the preamble should limit the body of the claim because the preamble had to be amended when the patent examiner found that the term "genomic sample" in the preamble was too vague and indefinite. The term "genomic sample," however, was found in both the preamble and the body of the text and was amended in both places to read "genomic DNA sample." (Decl. of Gottfried Ex. R 8 (dkt. no. 84-19).) Where the same language is found in both the preamble and body, the language in the preamble does not provide "essential

structure or steps" and is not "necessary to give life, meaning, and vitality to the claim." See Symantec, 522 F.3d at 1288; see also Avidyne Corp. v. L-3 Commc'ns Avionics Sys., Inc., 05-11098-GAO, 2008 WL 4849894, at *3 (D. Mass. Nov. 7, 2008) (finding preamble does not limit claim in part because "the terms in the preamble are not more specific or illuminating than the body of the claim").

The language of the preamble—"method for detecting the presence or absence of a single nucleotide polymorphism (SNP) allele in a genomic DNA sample"—is not limiting.

2. Body of Claim 1: "a reduced complexity genome (RCG)"

The parties dispute the proper construction for the term "a reduced complexity genome." Affymetrix argues that it means "a portion of the genomic DNA sample comprising two or more DNA fragments and which encompasses less than the entire native genome." (Joint Claim Construction Chart 2 (dkt. no. 109-2).) MIT argues that it means "a reproducible fraction of a genomic DNA sample that (1) is composed of a plurality of DNA fragments, and (2) has a sufficiently lower level of complexity than the original genomic DNA sample such that it is suitable for analysis by a hybridization technique for the presence or absence of a SNP allele." (Id.) Neither party's construction is entirely persuasive.

First, MIT is correct that, as used in the claim, an RCG must be "reproducible." The patent specification explicitly states that an "'RCG' as used herein is a *reproducible* fraction of an isolated genome which is composed of a plurality of DNA fragments," Col. 12:19-21 (emphasis added), and that "[t]he genotyping methods of the invention are based on use of RCGs that can be *reproducibly* produced," Col. 11:28-29 (emphasis added).

The second question is whether an RCG should be construed to be "a . . . fraction of a genomic DNA sample that is composed of a plurality of DNA fragments" (as MIT urges) or "a

portion of the genomic DNA sample comprising two or more DNA fragments and which encompasses less than the entire native genome" (as Affymetrix urges). (Joint Claim Construction Chart 2 (dkt. no. 109-2).) MIT's construction is rooted firmly in the patent language. See Col. 12:19-21 (an RCG is "a . . . fraction of an isolated genome which is composed of a plurality of DNA fragments."). The actual substantive difference between the two formulations is minimal. A "fraction" is easily understood as "encompass[ing] less that the entire native genome." MIT's proposed construction is adequate in this respect.

The third question is whether the term "reduced complexity genome" should also be construed, as MIT proposes, to have "a sufficiently lower level of complexity than the original genomic DNA sample such that it is suitable for analysis by a hybridization technique for the presence or absence of a SNP allele." (Joint Claim Construction Chart 2 (dkt. no. 109-2).) The suggested language is either redundant surplusage, in which case it is unnecessary, or it reads in an unclaimed limitation from the specification, in which case it is improper. See Ventana Med. Sys., Inc. v. Biogenex Labs., Inc., 473 F.3d 1173, 1181 (Fed. Cir. 2006).

As used in Claim 1, a "reduced complexity genome" means *a reproducible fraction of a genomic DNA sample that is composed of a plurality of DNA fragments*.

3. Body of Claim 1: "a randomly primed PCR-derived RCG"

The parties dispute the proper construction for "a randomly primed PCR-derived RCG."

Affymetrix argues that the term "randomly primed" should be construed to mean "primed with [polymerase chain reaction ("PCR")] primers that contain one or more arbitrarily selected nucleotide(s) such that the sequences amplified are not known in advance." (Joint Claim Construction Chart 2 (dkt. no. 109-2).) This is, for the most part, an acceptable construction of the words "randomly primed."

This construction is largely consistent with the language of the specification. Although the specification addresses several PCR methods, DOP-PCR appears to serve as the preferred embodiment, because it is discussed first and in somewhat greater detail than the other methods. It is clear that DOP-PCR as described in the specification employs primers that have "arbitrarily selected" nucleotide sequences. See, e.g., Col. 15:18-20 ("The 'TARGET' nucleotide sequence includes at least 5 arbitrarily selected nucleotide residues that are the same for each primer in the set."); Col. 15:48-51 ("The arbitrarily selected sequence is positioned at the 3' end of the primer. This sequence, although arbitrarily selected, is the same for each primer in a set of DOP-PCR primers."); Col. 16:10-14 ("Because DOP-PCR involves a randomly chosen sequence, the resultant PCR products are generated from genome sequences arbitrarily distributed throughout the genome and will generally not be clustered within specific sites of the genome."). The same is true of "arbitrarily primed" (or AP) PCR. See Col. 17:31-34 ("For AP-PCR, the primer sequence is arbitrary and is selected without knowledge of the sequence of the target nucleic acids to be amplified."). Similarly, the specification identifies a manner of performing an adapter-PCR that would have at least an element of randomness. See Col. 18:48 (referring to the use of a "random 3 base pair sequence"). It is not clear whether randomness is or can be employed in an IRS-PCR, also described in the specification, but that is not especially significant for the present question because, as Affymetrix points out, it is not necessary for the claim to cover every embodiment discussed in the specification. See Ventana, 473 F.3d at 1181.

The final portion of Affymetrix's proposed construction, stating that the nucleotides are selected "such that the sequences amplified are not known in advance," is unclear. It is a truism that a sequence to be selected randomly is unknown in advance of its selection, so if that is what

"known in advance" is intended to mean, it is unnecessary. If it is meant to describe something else, then it is unexplained. The phrase should be omitted.

The understanding of "randomly primed," as described above, can then be used in establishing the meaning of the full phrase "randomly primed PCR-derived RCG." The full phrase is construed to mean *a reduced complexity genome (RCG) that is generated by amplification of a set of DNA fragments by means of a polymerase chain reaction (PCR) that is primed with PCR primers that contain one or more arbitrarily selected nucleotides.*

4. Body of Claim 1: "analyzing the RCG directly by hybridization for the presence or absence of a SNP allele"

The question here is the meaning of "directly by hybridization." The claim history supports MIT's proposal. MIT added the language "directly by hybridization" to distinguish its work from a prior invention that would require an additional step in which the RCG would be broken into smaller segments before analysis. (See Decl. of Gottfried Ex. T 11-12 (dkt. no. 84-21).) In light of this, "directly" is fairly understood to mean "without intermediate steps."

Further, it is sensible to construe "hybridization" to mean "the process of hybridization," so as to include whatever ancillary steps or procedures are normally associated with that process, such as standard laboratory protocols.

Accordingly, as used in Claim 1, the term "analyzing the RCG directly by hybridization for the presence or absence of a SNP allele" means *the randomly-primed PCR-derived RCG is analyzed by a hybridization technique to detect the presence or absence of one or more SNP alleles without performing an intermediate step (such as locus-specific PCR) designed to further isolate a minor subset of the amplified DNA sequences. Standard laboratory steps that would ordinarily be performed in connection with hybridization of an RCG may be performed in accordance with general practice.*

B. Claim 2

Claim 2 of the '228 Patent reads:

> The method of claim 1, wherein the analysis comprises hybridizing a SNP-ASO and the RCG, wherein the SNP-ASO is complementary to one allele of a SNP, whereby the allele of the SNP is present in the genomic DNA sample if the SNP-ASO hybridizes with the RCG, and wherein the presence or absence of the SNP allele is used to characterize the genomic DNA sample.[3]

1. Body of Claim 2: "wherein the SNP-ASO is complementary to one allele of a SNP"

The parties dispute the proper construction for the phrase "wherein the SNP-[allele specific oligonucleotide ("ASO")] is complementary to one allele of a SNP."

MIT argues that the language should be construed to mean: "The SNP-ASO has a sequence of nucleotides that allow it to preferentially hybridize under appropriate conditions with the portion of the sequence of a DNA strand that contains that SNP allele." (Joint Claim Construction Chart 6 (dkt. no. 109-2).) This proposed construction is satisfactory except for two matters. First, the word "preferentially" is both ambiguous and unsupported by the specification and should not be included. Second, it should be clear that the complementarity of the SNP-ASO with the allele of interest must be such that the SNP-ASO will form a base pair with the SNP allele of interest.

As used in Claim 2, "the SNP-ASO is complementary to one allele of a SNP" means that *the SNP-ASO has a sequence of nucleotides that allow it to hybridize under appropriate conditions with the portion of the sequence of a DNA strand that contains that SNP allele, forming a base pair with that allele.*

[3] The original patent language has been corrected. (See Affymetrix Claim Construction Presentation 118; Joint Claim Construction Chart 6-7 (dkt. no. 109-2).)

2. *Body of Claim 2: "the SNP-ASO hybridizes with the RCG"* and *"the presence or absence of the SNP allele used to characterize the genomic DNA sample"*

These phrases need no construction. They should be understood according to their plain meaning to a person of skill in the art.

C. Claim 20 and Claim 21

Claim 20 of the '228 Patent reads: "The method of claim 2, wherein the SNP-ASO is composed of from about 10 to about 50 nucleotide residues." Col. 242:18-19. Claim 21 reads: "The method of claim 20, wherein the SNP-ASO is composed of from about 10 to about 25 nucleotide residues." Col. 242:20-21.

These terms can be given their plain meaning: "about" means *approximately*.

Whether the claims, so construed, are too indefinite is a matter for another day.

D. Claim 25

Claim 25 of the Patent reads: "The method of claim 1, wherein the method further comprises identifying a genotype of the genomic DNA sample, whereby the genotype is identified by the presence or absence of the allele of the SNP in the RCG." Col. 242:29-33.

Affymetrix proposes language that closely tracks the language of the claim and is consistent with the patent claim history. It essentially restates the plain meaning.

As used in Claim 25, "the genotype is identified by the presence or absence of the allele of the SNP in the RCG" means *a genotype is assigned to the genomic DNA sample by determining the presence or absence of the allele of the SNP in the RCG*.

E. Claim 38 and Claim 39

Claim 38 reads: "The method of claim 1, wherein the complexity of the genome is reduced by 50%." Col. 244:5-6. Claim 39 reads: "The method of claim 1, wherein the complexity of the genome is reduced by 95%." Col. 244:5-6.[4]

Affymetrix argues that the claim is that the genome must be reduced by *exactly* those amounts, while MIT argues that the genome must be reduced by *at least* those amounts. The patentee chose the words, and could easily have added "at least" in each formulation. Or perhaps the claim could refer to "*about* 50%" or "*about* 95%," as in claims 20 and 21. As written, the claims are specific to the stated percentages. They should be given their plain and precise meaning.

/s/ George A. O'Toole, Jr.
United States District Judge

[4] The original patent language has been corrected. (See Affymetrix Claim Construction Presentation 150; Joint Claim Construction Chart 9 (dkt. no. 109-2).)